ISRAEL JOURNAL OF MATHEMATICS **102** (1997), 189–197

ON EMBEDDING EXPANDERS INTO ℓ_p SPACES

BY

JIŘÍ MATOUŠEK*

Department of Applied Mathematics, Charles University
Malostranské nám. 25, 118 00 Praha 1, Czech Republic
e-mail: matousek@kam.mff.cuni.cz

ABSTRACT

In this note we show that the minimum distortion required to embed all n-point metric spaces into the Banach space ℓ_p is between $(c_1/p)\log n$ and $(c_2/p)\log n$, where $c_2 > c_1 > 0$ are absolute constants and $1 \le p < \log n$. The lower bound is obtained by a generalization of a method of Linial et al. [LLR95], by showing that constant-degree expanders (considered as metric spaces) cannot be embedded any better.

1. Introduction

Let M be a metric space with a metric ρ, let X be a normed space (whose norm will be denoted by $\|.\|$), and let $f\colon M \to X$ be a mapping. We say that f is a **D-embedding** (or a mapping with *distortion* at most D), $D \ge 1$ a real number, if we have

$$\frac{1}{D}\,\rho(x,y) \le \|f(x) - f(y)\| \le \rho(x,y)$$

for any two points $x, y \in M$. We say that M D-embeds into X if there exists a D-embedding** $f\colon M \to X$.

* Research supported by Czech Republic Grant GAČR 201/94/2167 and Charles University grants No. 351 and 361.

** A number of various terms besides the mentioned ones are used in the literature in this context; e.g., a D-embedding is also called a *D-isomorphism*, a *D-lipeomorphism*, etc.

Received January 26, 1996

The D-embeddability of finite metric spaces into various normed spaces was investigated in the context of the local Banach space theory ([Efn69a], [Efn69b], [JL84], [Bou85], [BMW86], [JLS87], [AR92], [Ma95]), and it seems that it can be of considerable interest also in more applied areas (see [LLR95]).

Let the symbol ℓ_p^n denote the n-dimensional real vector space equipped with the L_p-norm, given by $\|(x_1, x_2, \ldots, x_n)\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$ (for $1 \leq p < \infty$). Similarly ℓ_p denotes the space of countable sequences of real numbers with a finite L_p-norm. For a metric space M, let $D_p(M)$ be the minimum D such that M can be D-embedded into ℓ_p, and let $D_p(n)$ be the supremum of $D_p(M)$ over all n-point metric spaces M. Since any n-point subset of ℓ_p can be isometrically embedded into $\ell_p^{n(n-1)/2}$ (see e.g. [Fic88]), we can restrict our attention to embeddings into finite-dimensional ℓ_p spaces.

The most well-studied case is that of $p = 2$, where we are dealing with embeddings into the usual Euclidean space. It is easy to find a 4-point metric space which cannot be isometrically embedded into any Euclidean space, but it is not so easy to prove that $D_2(n) \to \infty$ for $n \to \infty$; this was probably first done by Enflo [Efn69b], whose proof yields $D_2(n) = \Omega(\sqrt{\log n})$ (his example is the cube $\{0,1\}^k$ with the L_1 metric, $n = 2^k$; see also [Efn69a]). Bourgain [Bou85] proved an upper bound $D_2(n) = O(\log n)$ and a lower bound $D_2(n) = \Omega(\log n / \log\log n)$. The lower bound is non-constructive, using random graphs and a counting argument. Linial et al. [LLR95] discovered another lower bound technique, which allowed them to show the asymptotically tight lower bound $D_2(n) = \Omega(\log n)$; their proof, unlike Bourgain's, yields an explicit metric space exhibiting the lower bound (one can say that their method slightly resembles Enflo's, with expander graphs replacing the L_1-cube).

The situation for other values of p has been understood less satisfactorily. Concerning upper bounds, Bourgain's embedding technique in fact proves $D_p(n) = O(\log n)$ for any p (with the constant of proportionality independent of p). Concerning lower bounds, the argument of Linial et al. [LLR95] shows $D_p(n) = \Omega(\log n)$ for any $p \in [1,2]$. For $p > 2$, however, the best known lower bound was apparently one following from the results of Bourgain et al. [BMW86], which is $c_\varepsilon (\log n)^{1/2-\varepsilon}$, with $\varepsilon > 0$ an arbitrarily small number and $c_\varepsilon > 0$ depending on ε. Here we obtain asymptotically tight bounds:

THEOREM 1:

(i) *There exist constants $c_1 > 0$ and n_0 such that for any $p \geq 1$ and any*

$n \geq n_0$ there exists an n-point metric space which D-embeds into ℓ_p only for $D \geq (c_1/p)\log n$.

(ii) *Any n-point metric space can be embedded into ℓ_p with distortion at most $(c_2/p)\log n$, where c_2 is a constant and $1 \leq p < \log n$.*

Part (i) is proved in section 2 by generalizing the method of Linial et al. [LLR95]. Part (ii) is proved in section 4 by modifying Bourgain's embedding method very slightly.

2. Expanders

Let $G = (V, E)$ be a (simple, unoriented) graph on the vertex set $V = \{1, 2, \ldots, n\}$. We assume that G is d-regular (every vertex has exactly d neighbors) with d a constant (while n is a variable attaining arbitrarily large values). The graph G is called an **expander** if there exists a constant $\Phi > 0$ (independent of n; Φ is called the **conductance** of G) such that for any subset $A \subseteq V$ with $|A| \leq n/2$ we have

$$|\{\{i,j\} \in E; i \in A, j \in V \setminus A\}| \geq \Phi|A|. \tag{1}$$

It can be shown that a random d-regular graph is an expander (with a suitable $\Phi = \Phi(d)$) with a positive probability. Sophisticated explicit constructions of expanders are also known — see e.g. [AS92] for background information and references.

If $x_1, x_2, \ldots, x_n$ are real numbers, their **median** is defined as a real number m such that $|\{i; x_i \leq m\}| \geq \lfloor n/2 \rfloor$ and $|\{i; x_i \geq m\}| \geq \lfloor n/2 \rfloor$ (thus, for n even, the median need not be determined uniquely). A basic property of expanders we use is the following (apparently due to Sinclair and Jerrum [JS88]; see e.g. Lovász [Lov93], Ex. 11.30):

LEMMA 2: *Let G be an expander with conductance $\Phi > 0$, let $x_1, x_2, \ldots, x_n$ be arbitrary real numbers, and let m be their median. Then*

$$\sum_{\{i,j\} \in E} |x_i - x_j| \geq \Phi \sum_{i \in V} |x_i - m|. \tag{2}$$

The proof of (2) has few lines. Inequality (1) can be viewed as a special case of (2), where the x_i's only attain values 0 and 1 (set $x_i = 1$ for $i \in A$ and $x_i = 0$ otherwise).

The key to the non-embeddability result of Linial et al. [LLR95] is a suitable generalization of (2) for the x_i's being elements of a Banach space instead of real numbers. Let $K = \binom{V}{2}$ denote the set of edges of the complete graph with vertex set V. A simple consequence of (2) is the following:

$$\sum_{\{i,j\}\in E} |x_i - x_j| \geq \frac{\Phi}{n-1} \sum_{\{i,j\}\in K} |x_i - x_j|. \tag{3}$$

Indeed, since (3) is clearly translation-invariant (it is not changed by adding the same real number to each x_i), we may assume that 0 is a median of the x_i's. Then we have $\sum_{\{i,j\}\in K} |x_i - x_j| \leq \sum_{\{i,j\}\in K}(|x_i| + |x_j|) = (n-1)\sum_{i\in V} |x_i|$, so (3) follows from (2).

Linial et al. [LLR95] essentially observe (although they formulate it in a somewhat different manner) that (3) holds also for any $x_1, x_2, \ldots, x_n \in \ell_1$:

$$\sum_{\{i,j\}\in E} \|x_i - x_j\|_1 \geq \frac{\Phi}{n-1} \sum_{\{i,j\}\in K} \|x_i - x_j\|_1. \tag{4}$$

This is an immediate consequence of (3): write each $x_i \in \ell_1$ in coordinates, use (3) for each coordinate separately, and sum the resulting inequalities. Since any finite subset of any ℓ_p with $1 \leq p \leq 2$ can be isometrically embedded into ℓ_1 [BDK66], an analog of (4) holds also in these ℓ_p's, in particular in a Euclidean space.

For $p > 2$, we need an "L_p-analog" of (3):

PROPOSITION 3: *Let G be a d-regular expander with conductance Φ, let $x_1, x_2, \ldots, x_n$ be real numbers with median m, and let $p > 1$. Then we have*

$$\begin{aligned}\sum_{\{i,j\}\in E} |x_i - x_j|^p &\geq \frac{(\Phi/2p)^p}{d^{p-1}} \sum_{i\in V} |x_i - m|^p \\ &\geq \frac{(\Phi/4p)^p}{d^{p-1}} \frac{1}{n-1} \sum_{\{i,j\}\in K} |x_i - x_j|^p.\end{aligned} \tag{5}$$

The version of this result for $p = 2$ was essentially proved by Sinclair and Jerrum [JS88] (the method goes back to Alon [Alo86]); the result for a general p doesn't seem to be known. We prove it in section 3 below by generalizing the known proof of the $p = 2$ case (following the presentation of Lovász [Lov93], Ex. 11.32).

As a consequence we get that for any $x_1, x_2, \ldots, x_n \in \ell_p$,

$$\sum_{\{i,j\}\in E} \|x_i - x_j\|_p^p \geq \left(\frac{c_0}{p}\right)^p \frac{1}{n-1} \sum_{\{i,j\}\in K} \|x_i - x_j\|_p^p, \tag{6}$$

with $c_0 = \Phi/4d^{1-1/p}$ a positive constant (this again follows by applying (5) to each coordinate and summing up).

To derive Theorem 1(i) from (6), we consider the expander G as an n-point metric space, with the metric ρ given by the usual graph-theoretic distance of vertices. Let us consider the ratio

$$R_\rho = \frac{\left(\frac{1}{|E|}\sum_{\{i,j\}\in E} \rho(i,j)^p\right)^{1/p}}{\left(\frac{1}{|K|}\sum_{\{i,j\}\in K} \rho(i,j)^p\right)^{1/p}}.$$

The numerator is the pth degree average of the edge length, which is 1 by definition. The denominator is the pth degree average of the distance of two vertices of G. Since G is d-regular, at most $1 + d + \cdots + d^k \leq 2d^k$ vertices of G have distance at most k from a given vertex, and from this one can see that at least a fixed fraction of the pairs $\{i,j\} \in K$ satisfies $\rho(i,j) \geq \log_d(n/4)$ (say). Therefore $R_\rho = O(1/\log n)$.

Next, suppose that $f : V \to \ell_p$ is a D-embedding, and let σ be the metric on V given by $\sigma(i,j) = \|f(i) - f(j)\|_p$. Define the ratio R_σ analogously to R_ρ. Then inequality (6) shows that $R_\sigma \geq c'/p$, with $c' = \Phi/4d$ a positive constant (we use $|E| = dn/2$ and $|K| = n(n-1)/2$). On the other hand, if f is a D-embedding, we should have $R_\sigma \leq DR_\rho$, and hence $D \geq (c_1/p)\log n$ as claimed in Theorem 1.
∎

3. A p-inequality for expanders

For a real number $x < 0$, let x^p stand for $-(-x)^p$. First we note the following estimate:

LEMMA 4: *For any real numbers a, b and any $p \geq 1$, we have*

$$|a^p - b^p| \leq p|a-b|\left(|a|^{p-1} + |b|^{p-1}\right).$$

Proof: By symmetry, we may assume $a \geq |b| > 0$, and by re-scaling we may suppose $a = 1$, so it is enough to show $1 - b^p \leq p(1-b)$ for any $b \in (-1, 1)$.

Finally, by writing $x = 1 - b$ we pass to $(1-x)^p \geq 1 - px$ which is a well-known (Bernoulli's) inequality. ∎

Proof of Proposition 3: Let p, $x_1, x_2, \ldots, x_n$, and m be as in the Proposition; we may assume $m = 0$. For the sake of brevity, put

$$S = \sum_{\{i,j\}\in E} |x_i - x_j|^p, \quad T = \sum_{i\in V} |x_i|^p.$$

By (2) applied to the numbers $x_1^p, \ldots, x_n^p$ we have

$$\Phi T \leq \sum_{\{i,j\}\in E} |x_i^p - x_j^p|.$$

Using Lemma 4, we further get that the right-hand side is at most $p\sum_{\{i,j\}\in E} u_{ij}v_{ij}$, where $u_{ij} = |x_i - x_j|$, $v_{ij} = |x_i|^{p-1} + |x_j|^{p-1}$. By Hölder's inequality, we have

$$\sum_E u_{ij}v_{ij} \leq \left(\sum_E u_{ij}^p\right)^{1/p} \left(\sum_E v_{ij}^q\right)^{1/q}$$

where $q = p/(p-1)$. We note that $\sum_E u_{ij}^p = S$, and we estimate $v_{ij}^q = \left(|x_i|^{p-1} + |x_j|^{p-1}\right)^q \leq 2^q\left(|x_i|^{(p-1)q} + |x_j|^{(p-1)q}\right) = 2^q\left(|x_i|^p + |x_j|^p\right)$; hence, using the d-regularity of G, we get

$$\left(\sum_E v_{ij}^q\right)^{1/q} \leq 2\left(\sum_E \left(|x_i|^p + |x_j|^p\right)\right)^{1/q} = 2d^{1/q}\left(\sum_{i\in V}|x_i|^p\right)^{1/q} = 2d^{1/q}T^{1/q}.$$

Combining the whole chain of inequalities yields $\Phi T \leq pS^{1/p}2d^{1/q}T^{1/q}$, hence $S \geq (\Phi/2p)^p d^{-(p-1)}T$, which is the first inequality in Proposition 3. The second inequality follows by estimating $|x_i - x_j|^p \leq 2^p\left(|x_i|^p + |x_j|^p\right)$. ∎

4. An upper bound

Here we prove part (ii) of Theorem 1. We use the method invented by Bourgain [Bou85]. This method has been used, with small modifications, in a number of other papers ([JLS87], [Ma91], [LLR95], [Ma95]), and the author of the present note finds it already somewhat embarrassing to repeat essentially the same thing here again; so the presentation is somewhat sketchy. On the other hand, it is interesting that one can get a tight upper bound also in terms of p.

The proof is based on the following lemma:

LEMMA 5: *Let M be an n-point metric space with a metric ρ. Let x, y be two distinct points of M, and let $s \geq 2$ be a parameter. Then there exist real numbers $\Delta_1, \Delta_2, \ldots, \Delta_t \geq 0$ with $\Delta_1 + \cdots + \Delta_t = \rho(x,y)/4$, where $t = \lfloor \log_s n \rfloor + 1$, and such that the following holds for each $i = 1, 2, \ldots, t$: if $A_i \subseteq M$ is a randomly chosen subset of X, with each point of X included in A_i independently with probability $1/s^i$, then the probability P_i of the event "$|\rho(x, A_i) - \rho(y, A_i)| \geq \Delta_i$" satisfies $P_i \geq 1/8s$.*

Proof sketch: As shown in [LLR95] (or [Bou85] with a slightly different formulation)*, the numbers Δ_i can be chosen in such a way that P_i is at least the probability that $A_i \cap S_1 = \emptyset$ and at the same time $A_i \cap S_2 \neq \emptyset$, where S_1 is a certain subset of X of size $< s^i$ and S_2 is another subset of X, disjoint from S_1, of cardinality $\geq s^{i-1}$. A detailed calculation showing that the latter probability is at least $1/8s$ is given in [Ma95, Lemma 4.1]. ∎

Proof of Theorem 1(ii): Let (M, ρ) be a given n-point metric space. Fix $s = 2^p$, $t = \lfloor \log_s n \rfloor + 1$, and for each $i = 1, 2, \ldots, t$ choose r independent random subsets $A_{i1}, \ldots, A_{ir}$, each $A_{ij} \subseteq X$ being chosen as the A_i in Lemma 5 (i.e. each point included with probability $1/s^i$). If r is chosen sufficiently large, by Lemma 5 we may assume that *for each* $x, y \in X$ and each $i = 1, 2, \ldots, t$, the inequality $|\rho(x, A_{ij}) - \rho(y, A_{ij})| \geq \Delta_i$ holds for at least $r/16s$ indices j, where Δ_i depends on x, y and is as in the Lemma (one can take $r = \text{const}.s \log n$, as can be shown using a suitable version of the Chernoff inequality). We fix such a collection of the A_{ij} and define a mapping $f: M \to \ell_p^{tr}$: if the coordinates in ℓ_p^{tr} are indexed the same way as the sets A_{ij}, we define f componentwise by $f(x)_{ij} = \rho(x, A_{ij})$.

Since $|\rho(x, A) - \rho(y, A)| \leq \rho(x, y)$ holds for any set A, we obtain

$$\|f(x) - f(y)\|_p \leq t^{1/p} r^{1/p} \rho(x, y).$$

On the other hand, we have

$$\|f(x) - f(y)\|_p = \left(\sum_{i=1}^{t} \sum_{j=1}^{r} |\rho(x, A_{ij}) - \rho(y, A_{ij})|^p \right)^{1/p} \geq \left(\sum_{i=1}^{t} \frac{r}{16s} \Delta_i^p \right)^{1/p};$$

using Hölder's inequality, we get $\sum_{i=1}^{t} \Delta_i^p \geq \left(\sum_i \Delta_i \right)^p / t^{p-1} = (\rho(x,y)/4)^p / t^{p-1}$,

* The proofs in [LLR95] and [Bou85] argue for the $s = 2$ case, but the generalization to an arbitrary s is entirely straightforward.

hence $\|f(x) - f(y)\|_p \geq \frac{1}{64}(r/s)^{1/p}t^{-(p-1)/p}\rho(x,y)$. Thus after an appropriate scaling, f is a D-embedding with $D = O(t) = O(\log n/p)$. ∎

References

[Alo86] N. Alon, *Eigenvalues and expanders,* Combinatorica **6** (1986), 83–96.

[AS92] N. Alon and J. Spencer, *The Probabilistic Method,* Wiley, New York, 1992.

[AR92] J. Arias-de-Reyna and L. Rodríguez-Piazza, *Finite metric spaces needing high dimension for Lipschitz embeddings in Banach spaces,* Israel Journal of Mathematics **79** (1992), 103–111.

[Bou85] J. Bourgain, *On Lipschitz embedding of finite metric spaces in Hilbert space,* Israel Journal of Mathematics **52** (1985), 46–52.

[BDK66] J. Bretagnolle, D. Dacunha-Castelle and J. L. Krivine, *Lois stables et espaces L^p,* Annales de l'Institut Henri Poincaré. Probabilités et Statistiques, Sect. **B2** (1966), 231–259.

[BMW86] J. Bourgain, V. Milman and H. Wolfson, *On type of metric spaces,* Transactions of the American Mathematical Society **294** (1986), 295–317.

[Enf69a] P. Enflo, *On a problem of Smirnov,* Arkiv Matematik **8** (1969), 107–109.

[Enf69b] P. Enflo, *On the nonexistence of uniform homeomorphisms between L_p-spaces,* Arkiv Matematik **8** (1969), 103–105.

[Fic88] B. Fichet, *L_p-spaces in data analysis,* in *Classification and Related Methods of Data Analysis* (H. H. Bock, ed.), North-Holland, Amsterdam, 1988, pp. 439–444.

[JL84] W. Johnson and J. Lindenstrauss, *Extensions of Lipschitz maps into a Hilbert space,* Contemporary Mathematics **26** (Conference in Modern Analysis and Probability), American Mathematical Society, 1984, pp. 189–206.

[JLS87] W. Johnson, J. Lindenstrauss and G. Schechtman, *On Lipschitz embedding of finite metric spaces in low dimensional normed spaces,* in *Geometrical Aspects of Functional Analysis* (J. Lindenstrauss and V. D. Milman, eds.), Lecture Notes in Mathematics **1267**, Springer-Verlag, Berlin–Heidelberg, 1987, pp. 177–184.

[JS88] M. Jerrum and A. Sinclair, *Conductance and the rapid mixing property for Markov chains: The approximation of the permanent resolved,* in *Proceedings of the 20th ACM Symposium on Theory of Computing,* 1988, pp. 235–244.

[LLR95] N. Linial, E. London and Yu. Rabinovich, *The geometry of graphs and some of its algorithmic applications*, Combinatorica **15** (1995), 215–245.

[Lov93] L. Lovász, *Combinatorial Problems and Exercises* (2nd ed.), Akadémiai Kiadó, Budapest, 1993.

[Ma91] J. Matoušek, *Note on bi-Lipschitz embeddings into normed spaces*, Commentationes Mathematicae Universitatis Carolinae **33** (1992), 51–55.

[Ma95] J. Matoušek, *On the distortion required for embedding finite metric spaces into normed spaces*, Israel Journal of Mathematics **93** (1996), 333–344.